Analytical and Bioanalytical Chemistry

Electronic Supplementary Material

Targeted serum metabolite profiling and sequential metabolite ratio analysis for colorectal cancer progression monitoring

Jiangjiang Zhu, Danijel Djukovic, Lingli Deng, Haiwei Gu, Farhan Himmati, Mohammad Abu Zaid, Elena Gabriela Chiorean and Daniel Raftery

Supplemental Materials and Methods:

Chemicals and reagents: LC-MS grade acetonitrile, ammonium acetate, and acetic acid were purchased from Fisher Scientific (Pittsburgh, PA). Standard compounds corresponding to the measured metabolites were purchased from Sigma-Aldrich (Saint Louis, MO) or Fisher Scientific (Pittsburgh, PA), and a list of these compounds can be found in the Supplementary Table S1. Stable isotope-labeled tyrosine and lactate (L-tyrosine-$^{13}C_2$ and sodium-L-lactate-$^{13}C_3$) were purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA). The purities of non-labeled standards were >95-99%, whereas the purities of the two ^{13}C labeled compounds were >99%.

Sample preparation: Frozen samples were first thawed at room temperature for approximately 45 min, and 50 µL of each sample was protein precipitated using two rounds of cold methanol extraction (150 µL and 300 µL, respectively) at -20 °C. The resulting supernatant containing desired metabolites was collected into a new Eppendorf vial, dried using a Vacufuge Plus evaporator (Eppendorf, Hauppauge, NY), and then reconstituted in a 500 µL solution (40% water / 60% acetonitrile with 5 mM ammonium acetate and 0.2% acetic acid) containing 5.13 µM L-tyrosine-$^{13}C_2$ and 22.5 µM sodium-L-lactate-$^{13}C_3$. The two isotope-labeled internal standards were added to monitor system performance. The samples were filtered through 0.45 µm PVDF filters (Phenomenex, Torrance, CA) prior to LC-MS analysis. A pooled human serum sample was extracted using the same procedure as above. This sample was used as the quality control (QC) sample and was analyzed once every ten serum samples. All patient samples were randomized before LC-MS analysis.

LC-MS/MS system and conditions: The LC system consisted of two Agilent 1260 binary pumps, an Agilent 1260 auto-sampler, and an Agilent 1290 column compartment containing a

column-switching valve (Agilent Technologies, Santa Clara, CA). Two separate injections (10 µL for analysis using negative ionization mode and 2 µL for analysis using positive ionization mode) were made for each sample. Chromatographic separations were performed using hydrophilic interaction chromatography (HILIC) on two SeQuant ZIC-cHILIC columns (150 x 2.1 mm, 3.0 µm particle size, Merck KGaA, Darmstadt, Germany) connected in parallel. This setup facilitates high-throughput analysis as it allows one column to perform the separation while the other column is being reconditioned for the next sample injection. The reconstituted serum samples were gradient-eluted at 0.300 mL/min using solvents A (5 mM ammonium acetate in 90% water / 10% acetonitrile + 0.2% acetic acid) and B (5 mM ammonium acetate in 90% acetonitrile / 10% water + 0.2% acetic acid). The auto-sampler temperature was kept at 4 °C, the column compartment was set at 40 °C, and the separation time for each ionization mode was 20 min. The gradient conditions for both separations were identical and are briefly summarized as follows: 75% B isocratic for 2 min, 75% B to 30% B in 3 min, 30% B isocratic for 4 min, back to 75% B in 2 min, and then remaining at 75% B for 9 min.

The metabolite identities were confirmed by spiking the pooled serum sample used for method development with mixtures of standard compounds (each mixture contained five standard metabolites). The few metabolites that could not be well separated and had similar m/z values (<1 Da) were integrated as single peaks (e.g., malonic acid and 3-hydroxybutyric acid).

The mass spectrometer setting was optimized and described as follows. Briefly, after the chromatographic separation, MS ionization and data acquisition were performed using an AB Sciex QTrap 5500 mass spectrometer (AB Sciex, Toronto, ON, Canada) equipped with an electrospray ionization (ESI) source. The instrument was controlled by Analyst 1.5 software (AB Sciex). Targeted data acquisition was performed in multiple-reaction-monitoring (MRM) mode.

We monitored 105 and 57 MRM transitions in negative and positive mode, respectively (162 transitions in total). The source and collision gas was N_2 (99.999% purity). The ion source conditions in negative/positive mode were as follows: curtain gas (CUR) = 25 psi, collision gas (CAD) = high, ion spray voltage (IS) = -3.8/3.8 KV, temperature (TEM) = 500 °C, ion source gas 1 (GS1) = 50 psi, and ion source gas 2 (GS2) = 40 psi. The optimized MS conditions for each compound were optimized with chemical standards.

Supplemental data:

Table S1 List of targeted metabolites in this study (verified by chemical standards)

Glycine	Normetanephrin	Ribose-5-P
Trimethylamine-N-oxide	Histamine	Adenylosuccinate
Alanine	Pyruvate	D-Leucic acid
Aminoisobutyrate	Lactate	GDP
Choline	Acetoacetate	GTP
Dimethylglycine	Fumaric	DCDP
Serine	Succinate	Pyridoxal-5-P
Creatinine	Nicotinate	Gibberellin
Proline	Glutaric acid	Adipic acid
Valine	Malate	Maleic acid
Betaine	Hypoxanthine	Methylmalonate
Threonine	alpha-Ketoglutaric acid	DHAP
Taurine	Xanthine	Chenodeoxycholate
Creatine	PPA	G16BP
Hydroxyproline	Urate	F6P/F1P
Leucine/iso-Leucine	Homogentisate	Oxalic acid
Ornithine	PEP	Glyceraldehyde
Homocysteine	D-GA3P	Glycerate
Acetylcholine	Glycerol-3-P	N-Acetylglycine
Glutamine	Hyppuric acid	Guanidinoacetate
Glutamic acid	Glucose	Mevalonate
Methionine	4-Pyridoxic acid	Allantoin
Cystamine	2/3-Phosphoglyceric acid	Inositol
Histidine	Erythrose	Homovanilate
Carnitine	Cystathionine	Xanthurenate
Phenylalanine	G1P/G6P	Pentothenate
Arginine	Reduced glutatione	Biotin
Glucosamine	F16BP/F26BP	DCMP
Tyrosine	Sucrose	DUMP
Sorbitol	5-Formyl THF	Geranyl pyrophosphate
Epinephrine	Oxidized glutathione	DTMP
Tryptophan	gamma-Aminobutyrate	CMP
5-Hydroxytryptophan	Malonic acid/3HBA	Lactose
Uridine	Citraconic acid	cGMP
Phosphotyrosine	Adenine	AMP
Adenosine	Shikimic acid	IMP
Inosine	Aconitate	PGE

Guanosine	Citrulline	OMP
XMP	Citric acid	UDP
L-Kinurenine	Cystine	ADP
Lysine	Xanthosine	Folic acid
Cytosine	Uracil	DUTP
Homoserine	OH-Phenylpyruvate	ATP
Niacinamide	Glycochenodeoxycholate	Taurocholate
1-Methylhistamine	Glycocholate	Fructose
Asparagine	Dopamine	Aspartic acid
Salicylurate	Melatonin	Methylsuccinate
2'-Deoxyuridine	Orotate	Myristic acid
3-Hydroxykynurenine	Anthranilate	Margaric acid
Cytidine	Glucoronate	Linoleic acid
Pyroglutamic acid	Oxaloacetate	Linolenic acid
1-Methyladenosine	Propionate	Galactose
1-Methylguanosine	2-Aminoadipate	
N2,N2-Dimethylguanosine	Kynorenate	
Aminolevulinic acid	3-Nitro-tyrosine	

69

70

Table S2 Detailed patient and sample information

Patient	Gender	Age at Consent	BMI[a]	Diagnosis	Stage	# of blood draws	Disease Status [b] (follow the blood draw order)
1	M	29	30.4	Colon Cancer	Stage IV	4	DP, SD, DP, DP
2	M	32	28.1	Colon Cancer	Stage III	2	DP, DP
3	M	37	-	Colon Cancer	Stage IV	2	DP, DP
4	F	42	18.9	Colon Cancer	Stage IV	3	RD, DP, SD
5	F	42	-	Colon Cancer	Stage I/II	2	AD, CR
6	F	45	28.1	Colon Cancer	Stage IV	2	DP, DP
7	F	45	20.4	Colon Cancer	Stage IV	3	DP, DP, DP
8	F	46	27.5	Rectal Cancer	Stage IV	2	TR, CR
9	F	50	22.7	Colon Cancer	Stage IV	2	DP, DP
10	M	51	29.0	Colon Cancer	Stage III	2	AD, CR
11	F	55	25.6	Colon Cancer	Stage IV	4	LM, CR, CR, CR
12	M	55	-	Colon Cancer	Stage IV	2	DP, SD
13	M	65	24.4	Rectal Cancer	Stage III	2	CR, CR
14	F	66	-	Colon Cancer	Stage I/II	3	AD, CR, CR
15	F	66	24.0	Colon Cancer	Stage IV	3	DP, SD, DP
16	F	68	22.3	Rectal Cancer	Stage IV	3	DP, SD, DP
17	M	68	27.6	Colon Cancer	Stage IV	2	SD, SD
18	M	73	-	Colon Cancer	Stage IV	2	SD, SD
19	M	77	32.2	Colon Cancer	Stage III	2	SD, SD
20	F	86	25.4	Colon Cancer	Stage IV	2	DP, DP

[a] BMI information was not obtained for five patients.

[b] Disease status: DP, Disease Progression; AD, At Diagnosis; SD, Stable Disease; CR, Complete Remission; RD, Recurrent Disease; TR, Tumor Response; LM, Liver Metastasis. There were only three groups of disease status (CR, DP and SD) remaining after the ratio calculation.

Table S3 Summary of metabolites with high VIP scores (VIP>1.5) using sequential metabolite ratios in comparing DP vs. CR + SD

Metabolite	VIP	FC*	FDR
N2,N2-Dimethylguanosine	2.15	1.34	0.18
Citraconic Acid	2.04	1.58	0.29
1-Methylguanosine	2.04	1.25	0.29
Succinate	2.01	1.33	0.23
Adenine	2.01	1.11	0.37
Methylmalonate	1.84	1.31	0.31
3-Nitro-tyrosine	1.83	0.84	0.20
Malonic Acid/3HBA	1.77	5.99	0.14
G16BP	1.76	1.05	0.17
Urate	1.76	1.15	0.33
Aconitate	1.73	1.45	0.31
Homogentisate	1.66	1.19	0.14
Methylsuccinate	1.61	0.80	0.41
1-Methyladenosine	1.61	1.15	0.17
Cystathionine	1.60	0.65	0.20
Linolenic Acid	1.58	1.77	0.40
Cytidine	1.57	1.39	0.36
Pyruvate	1.57	0.64	0.33
Alanine	1.55	0.80	0.25
gamma-Aminobutyrate	1.53	0.82	0.44

*Fold change represents the average metabolite ratio for disease progression samples compared to samples from complete remission and stable disease.

Table S4 Summary of PLS-DA model performance using different numbers of sequential metabolite ratios for the differentiation of DP vs. CR + SD

PLS-DA Models	Metabolites selection Threshold	# of Metabolites used in the model[a]	AUROC	Sensitivity	Specificity
DP vs. CR+SD (Metabolites only models)	VIP>1.5	20	0.92	0.92	0.88
	VIP>1.8	7	0.90	0.83	0.94
	VIP>2	5	0.91	0.83	0.94
DP vs. CR+SD (Metabolites + CEA models)	VIP>1.5	20	0.92	0.92	0.88
	VIP>1.8	7	0.89	0.83	0.94
	VIP>2	5	0.91	0.83	0.94

[a] See Table S3 for metabolites and their corresponding VIP scores

Table S5 Summary of metabolites with *p*-value <0.05 using sequential metabolite ratios in comparing DP vs. CR + SD in stage IV patients

Metabolites	*p*-value	FC*	FDR
1-Methylguanosine	7.1E-03	1.27	0.18
N2,N2-Dimethylguanosine	1.3E-02	1.33	0.18
Adenine	1.5E-02	1.12	0.31
Succinate	1.8E-02	1.33	0.25
Pyruvate	3.0E-02	0.62	0.27
Methylmalonate	3.0E-02	1.28	0.25
Homogentisate	3.0E-02	1.22	0.14
Urate	3.0E-02	1.18	0.25
Citraconic Acid	4.9E-02	1.47	0.25

*Fold change represents the average metabolite ratio for disease progression samples compared to samples from other groups.

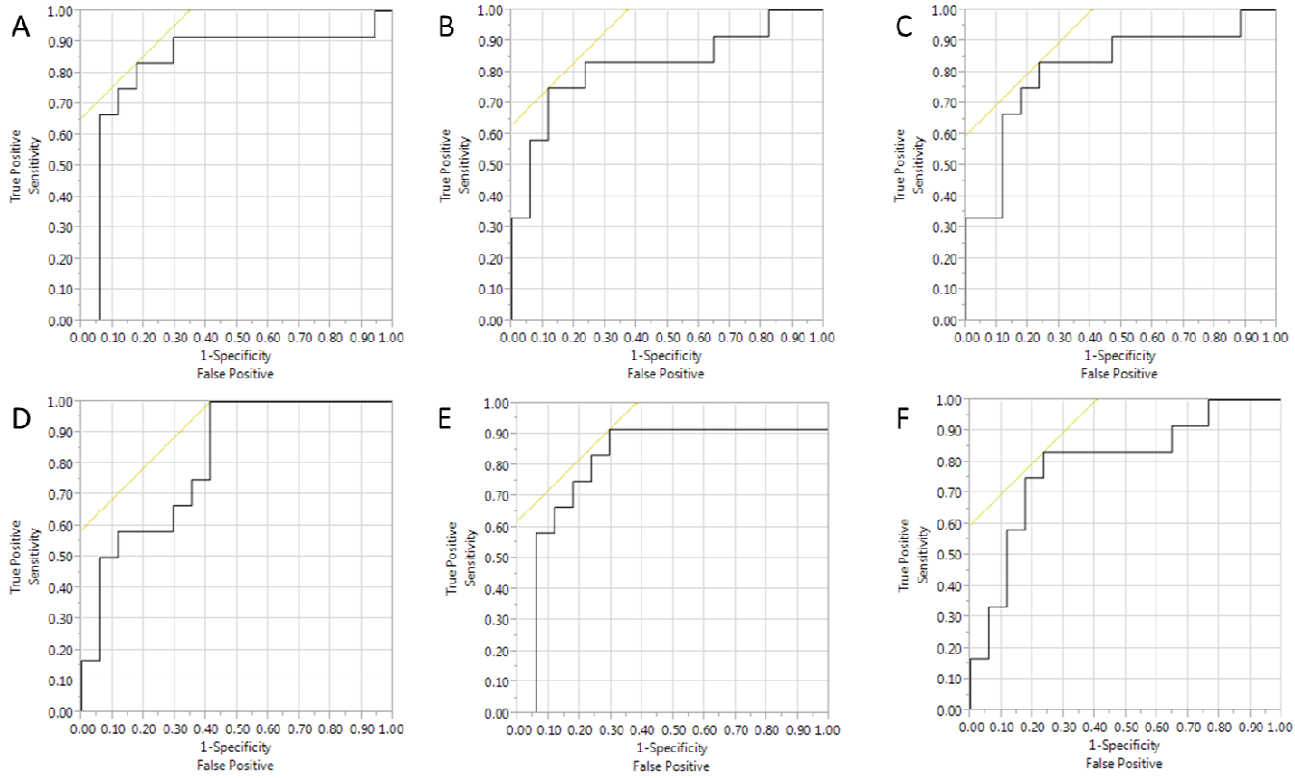


Fig. S1 Individual ROC curves for the top six metabolites with *p*-value<0.01 comparing DP with CR and SD using sequential metabolite ratios: (A) succinate, AUROC=0.83; (B) N2,N2-dimethylguanosine, AUROC=0.82; (C) citraconic acid, AUROC=0.81; (D) adenine, AUROC=0.81; (E) methylmalonate, AUROC=0.81; and (F) 1-methylguanosine, AUROC=0.79

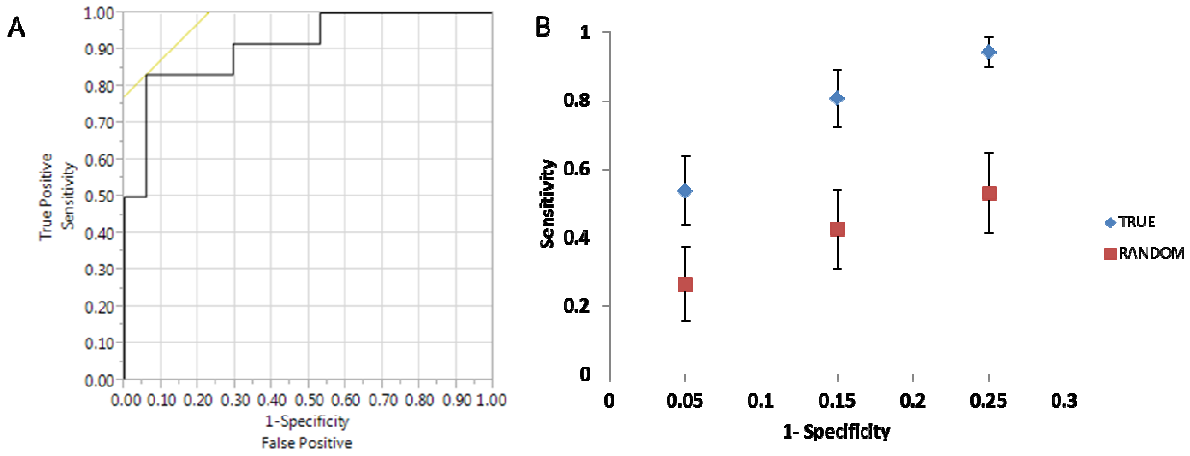


Figure S2 (A)ROC of PLS-DA model using five metabolites (with VIP>2) and CEA ratios for DP vs. CR + SD: AUROC= 0.912 (increased from 0.907, see Figure S3); sensitivity= 0.83; specificity= 0.94, FDR=0.09. (B) Monte Carlo cross validation (MCCV) PLS-DA results using the same metabolites: True, true class models; Random, random permutation model. The testing specificities were 0.95, 0.85, and 0.75. Error bars showing the standard deviation of 100 round of MCCV results

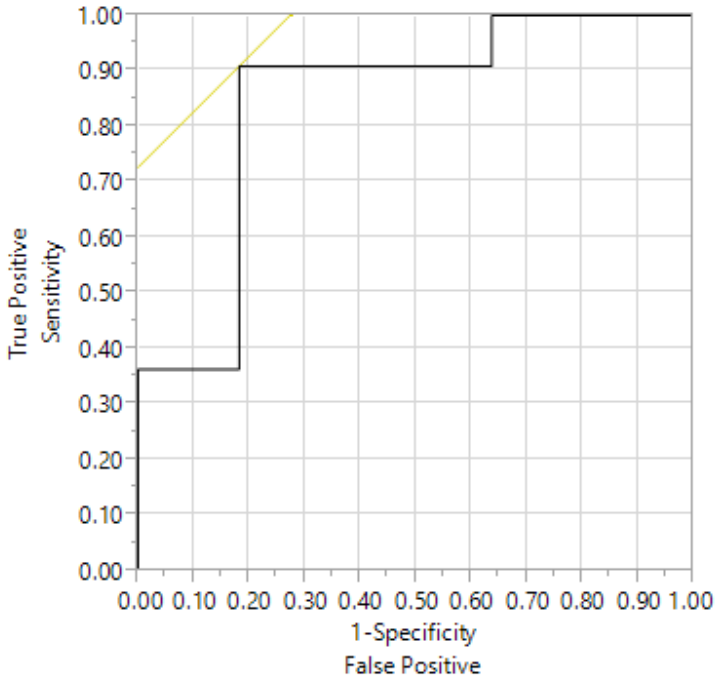
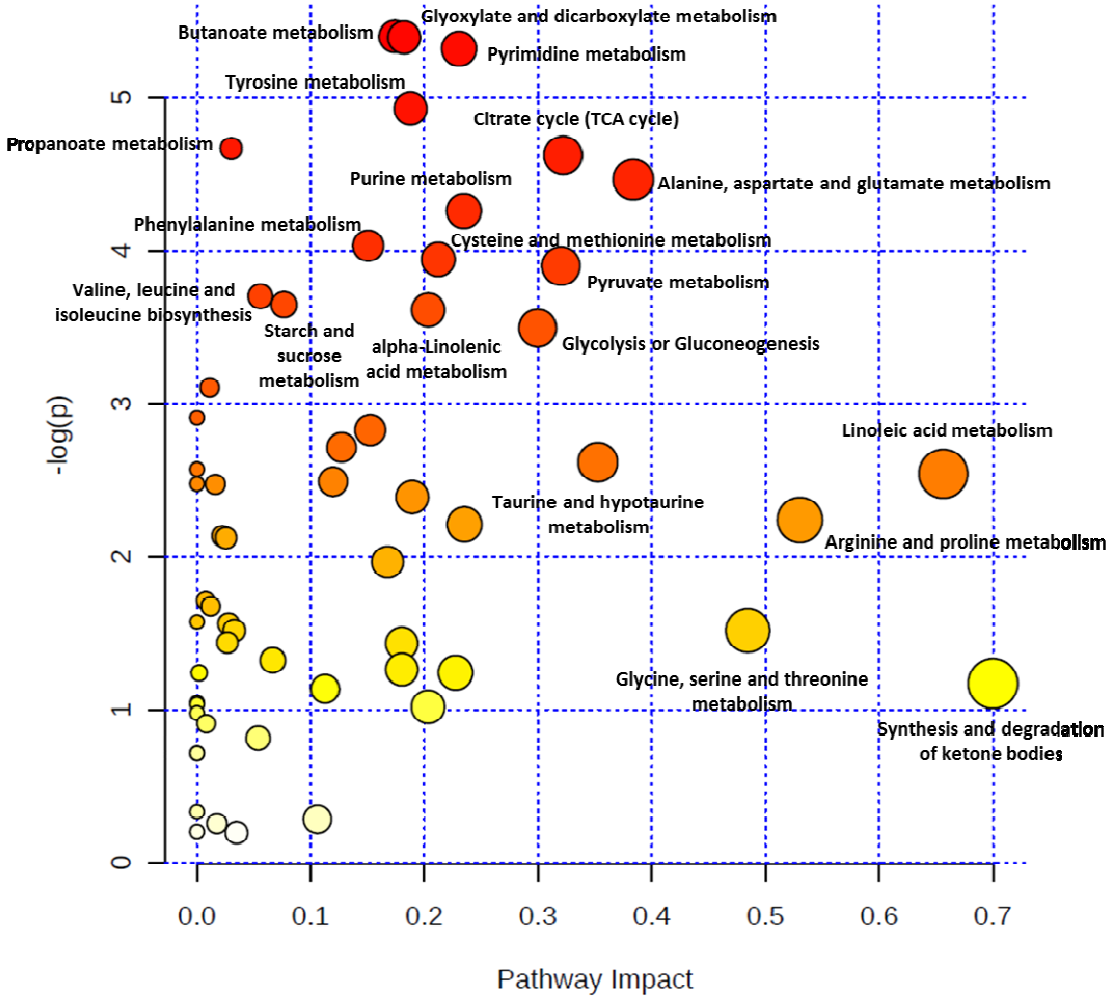


Fig. S3 ROC of PLS-DA model using five metabolites (with VIP>2) and CEA ratios for DP vs. CR + SD in stage IV patients: AUROC= 0.84, sensitivity= 0.91, specificity= 0.82, FDR=0.17



127

128 **Fig. S4** A metabolome view showing all impacted metabolic pathways in this study analyzed
129 using MetaboAnalyst(2.0), and using both scores from enrichment analysis (y axis) and from
130 topology analysis (x axis). Due to space restriction only –log(p) > 3 and pathway impact score >
131 0.3 are labeled

132